Biochemical and Biological Characterization of *D. Discoideum* TLP4, a Non-Canonical 3'-5' Polymerase

A Senior Thesis

Presented in partial fulfillment of the requirements for graduation

with research distinction in the undergraduate colleges of The Ohio State University

by

Jennie E. Hazen

The Ohio State University

April 2016

Project Advisor: Dr. Jane E. Jackman, Department of Chemistry and Biochemistry

Abstract

Post-transcriptional modification is an important step in the generation of many RNAs. For example tRNA^{His} requires the addition of a 5'-guanosine (G$_{-1}$), a reaction that is catalyzed by the tRNA^{His} guanylyltransferase (Thg1). Thg1 and related enzymes, Thg1-like proteins (TLPs), make up the Thg1/TLP enzyme family which spans all three domains of life and whose family members act as unusual RNA polymerases. While other polymerases act to extend a growing polynucleotide chain in the 5'-3' direction, Thg1/TLP family members act as non-canonical 3'-5' polymerases by adding one or more base-paired nucleotides to the 5'-ends of RNA. Despite their shared chemistry, several features distinguish these two types of enzymes. While TLPs are capable of adding nucleotides to a variety of substrates, Thg1 enzymes selectively recognize a single tRNA (tRNA^{His}) for G$_{-1}$ addition. Neither the molecular basis for Thg1's selectivity nor the basis for the observed differences between Thg1 and TLP functions are fully understood. In addition, while the biological function of Thg1 in tRNA^{His} maturation is well established, the role(s) of most TLPs remain poorly understood. This project aims to address both of those questions through the study of an essential TLP found in the slime mold *Dictyostelium discoideum*, DdiTLP4. DdiTLP4 exhibits distinct substrate specificity from Thg1, catalyzing nucleotide addition to non-tRNA substrates such as 5S rRNA and class I noncoding RNA (ncRNA) in vitro, although the biological substrates for the enzyme in vivo remain unknown. Site-directed mutagenesis and enzyme assays with purified proteins have been used to investigate the unique features of DdiTLP4 that allow it to recognize ncRNA, and to assess its activity with model stem-loop substrates, in order to compare its activities to the prototypical Thg1 enzyme. Genetic methods have also been used to determine the effects of depletion of DdiTLP4 on *D. discoideum*

development, to assess the effects of this unusual enzyme in biology.

Background

An organism's genotype manifests itself into an outward phenotype through a flow of information following a process called the "central dogma" that is shared amongst almost all organisms across all three domains of life. At the most basic level, genetic information is encoded into DNA. DNA is composed of two strands of deoxyribose sugars, which are each N-glycosidic linked to purine and pyrimidine bases, and linked to each other via 5'-3' phosphodiester bonds. To express its genetic material, the organism first undergoes a process called transcription, where the information contained in DNA strands are transcribed into RNA, the ribose-based complement to DNA, using the DNA molecule as a template for transcription. Both DNA and RNA convey their information through the sequence of their nucleotide bases, making up the specific region being transcribed. The information in the RNA is then translated into a protein molecule via the process of translation. The protein goes on to perform its encoded function in the organism, thus manifesting its encoded information into an outward phenotype. While some organisms such as retroviruses are capable of using their RNA as templates to make DNA, for the most part the pathway of DNA → RNA → Protein is highly conserved.

The process of transcription is carried out by enzymes called RNA polymerases, which interact with single strands of unwound DNA and use these strands as templates to produce pre-messenger RNA (pre-mRNA). Canonical RNA polymerases transcribe pre-messenger RNA in the 5' to 3' direction. In 5'-3' polymerization, the OH group of the 3' nucleotide of the growing polynucleotide chain attacks the α-phosphate of the free nucleotide, releasing pyrophosphate

and forming a phosphodiester bond. The generated pre-mRNA then undergoes further processing into mature mRNA, which is then translated into protein.

The process of translation is mediated by numerous RNAs, including small RNA molecules called transfer RNAs (tRNAs). These molecules help to decode the message in the messenger RNA (mRNA) and physically incorporate the correct amino acid into the growing polypeptide chain. Each tRNA molecule includes a three base-pair anticodon loop in its structure, which can bind to a complementary 3 base-pair codon sequence in the mRNA [figure 1]. Upon binding to the mRNA, the tRNA releases the amino acid bound to the CCA sequence on its 3' acceptor stem, and the ribosome incorporates the amino acid into the growing polypeptide chain. Each anticodon (and thus, each tRNA) corresponds to a specific amino acid, which is incorporated onto the tRNA in a process called aminoacylation by an aminoacyl-tRNA synthetase (aa-RS) enzyme. This step is a highly regulated process, as the proper aminoacylation of tRNAs is critical to ensure translational fidelity and the overall integrity of the protein. Incorporating the incorrect amino acid into a growing polypeptide chain could result in a functionally altered protein or in a misfolded, nonfunctional protein.

Unique nucleotide sequences called identity elements are located on specific regions of the tRNA molecule such as the anticodon, the acceptor stem, or both. The identity elements allow the tRNA to be recognized and aminoacylated by the proper aa-RS enzyme. For example, the identity element of tRNA histidine (tRNA^{His}) consists of a guanosine residue across from an adenine, at the 5' end of the tRNA's acceptor stem [figure 2]. While G_{-1} is encoded in the genomes of prokaryotic organisms, it needs to be post-transcriptionally added to the 5' end of tRNA^{His} in eukaryotes. This reaction is catalyzed by the RNA-editing polymerase tRNA^{His}

guanylyltransferase (Thg1) [1].

Thg1 and related enzymes, Thg1-like proteins (TLPs), make up the Thg1/TLP enzyme family which spans all three domains of life and whose family members act as unusual RNA polymerases. While other polymerases act to extend a growing polynucleotide chain in the 5'-3' direction, Thg1/TLP family members act as non-canonical 3'-5' polymerases by extending the 5' ends of RNA. In 3'-5' polymerization, the OH group of the free nucleotide attacks the α-phosphate of the 5' nucleotide on the growing polynucleotide chain, forming a new phosphodiester bond and releasing a pyrophosphate molecule. Thg1 is a 3'-5- polymerase capable of carrying out such a reaction. *In vivo,* it is capable of adding a non-Watson Crick paired G_{-1} to the 5' end of tRNA^{His}. Thg1 and TLPs add nucleotides to the 5' ends of RNAs through a conserved three-step mechanism. The activation step occurs first, where the enzyme adenylates the monophosphorylated 5' end of a substrate. The second step is the guanylyltransferase step itself, during which time the 3' OH- group of the incoming guanosine attacks the activated 5' end of the tRNA^{His}-AMP intermediate, incorporating itself into the RNA molecule and releasing the AMP molecule. The third and final step is a pyrophosphatase step that removes the pyrophosphate from the newly added G_{-1} [figure 3].

Despite their shared mechanism, several features distinguish Thg1 from the TLPs. Thg1 selectively recognizes a single tRNA substrate, tRNA^{His}, for a non-Watson Crick base paired nucleotide addition, G_{-1} across from A_{73}, *in vivo*. Other TLPs, however, are capable of recognizing and editing a wide variety of noncoding RNA substrates both *in vivo* and *in vitro,* and prefer to polymerize Watson-Crick base paired nucleotides [11]. The molecular bases for Thg1's narrow selectivity and the broader selectivity of the various TLPs are poorly understood. Additionally,

while the biological function of Thg1's G_{-1} addition activity is well understood, the biological functions of most TLP enzymes have yet to be elucidated.

This project aims to address these questions in order to gain a better understanding of the similarities and differences between Thg1 and the various TLPs through the study of an essential TLP, *Dictyostelium discoideum* TLP4 (DdiTLP4). The fourth TLP found in the slime mold *Dictyostelium discoideum*, DdiTLP4 is an essential enzyme that exhibits a very different substrate specificity from Thg1. It is capable of adding nucleotides to the truncated 5' ends of mitochondrial tRNAs *in vitro*, and it is also capable of adding nucleotides to non-tRNA substrates as well [2]. One of these potential substrates is the organism's 5s ribosomal RNA (5S rRNA), and another potential substrate is a small non-coding RNA called a class I non coding RNA unique to *D. discoideum* (ncRNA) [figure 4]. DdiTLP4 is capable of catalyzing nucleotide addition to both of these substrates *in vitro*, and ther is some evidence showing that it may be capable of acting on these substrates *in vivo* as well [2]. The function of the ncRNA is not yet known, but there is evidence that it may be involved in regulating the organism's development somehow, as these RNAs are down regulated during *D. discoideum*'s development [3].

Site-directed mutagenesis and enzyme assays with purified *S. cerevisiae* Thg1 (ScThg1), *Bacillus thuringiensis* TLP (BtTLP), and DdiTLP4 proteins have been used to investigate the unique structural features of DdiTLP4 that allow it to recognize ncRNA. Additionally, the kinetic properties of DdiTLP4 were measured using minimal RNA stem-loop substrates to compare its activity levels to those of TLPs from other domains of life as well as to those of the prototypical ScThg1 enzyme. Genetic methods have also been used to determine the effects of depletion of DdiTLP4, another *D. discoideum* TLP called DdiTLP3, which is capable of editing and repairing

mitochondrial tRNAs *in vivo* and some cytosolic RNAs *in vitro*, and Thg1 on *D. discoideum* development to assess the effects of this unusual enzyme in biology [3].

Chapter 1: *In Vivo* Characterization of DdiTLP4

Investigating the effect of TLP knockdown in *D. Discoideum* development

Dictyostelium discoideum expresses four TLPs, among which include DdiThg1, a homolog of human and yeast Thg1, DdiTLP2, DdiTLP3, and DdiTLP4 [4]. *D. discoideum* undergoes a unique life cycle characterized by two distinct phases of development: the unicellular vegetative phase and the multicellular social phase. The organism initially starts out as a single cellular haploid organism that can reproduce both asexually and sexually, and feeds on bacteria living in the soil. However, when confronted with a lack of nutrients the single-celled organisms go into a starvation mode, and enter the social cycle [5].

The single-celled amoeba begin to move as they undergo chemotaxis towards sources of cAMP, where they aggregate with one another and form a migrating "slug" that can move towards potential food sources at higher speeds than they could as individual amoeba. In the absence of any food source, the slug will then enter a culmination phase where it changes morphologically to form a stalk that supports fruiting body, which contains spores. The amoeba making up the slug will differentiate during this morphological process, with the cells at the anterior end of the slug dying and pushing upwards to form a stalk, and the cells at the posterior end forming a fruiting body. When nutrients become available, the fruiting body releases its spores, which were generated through both asexual and sexual reproduction. The spores then hatch into new *D. discoideum* amoeba and the cycle begins again [figure 5] [5].

In an effort to learn more about the *in vivo* functions of amoeboid Thg1/TLPs, and because previous data suggests that one of DdiTLP4's potential *in vivo* substrates may be involved in the organism's development, we used a TET-off RNAi knockdown model system to deplete DdiThg1, DdiTLP3, and DdiTLP4 expression levels during each of the two stages of *D. discoideum's* life cycle and look for any developmental phenotypes. Small-hairpin RNA encoding sequences complementary to the message of the gene of interest were fused to a tetracycline responsive promoter element (TRE) and then transformed into the *D. discoideum*. The TRE is composed of a promoter element fused to a Tet operator (TetO) sequence. The TRE activates transcription of the gene under its control when it interacts with the tetracycline transactivator (tTA) protein. In the presence of tetracycline, tTA preferentially binds to the antibiotic and the shRNA genes under the TRE's control cannot be transcribed. However, once tetracycline is removed from the media the tTA is free to bind to the TET promoter. The shRNA is transcribed, taking on the form of a small hairpin with a loop connected to a double stranded RNA molecule, and exported into the nucleus where it is cleaved by the slime mold's endogenous DICER complex into small double-stranded siRNA . The strands of siRNA are then separated and incorporated into RISC complexes, where they are directed to the gene whose sequence they are complementary to, binding to and degrading the gene's message [figure 6]. RNA levels for the different TLPs were not quantified for this experiment, but experiments using this TET-off system have reported RNA depletion levels of 40-50% [2]. Enzyme levels were depleted in the vegetative phase by removing tetracycline from two days prior to forcing the organisms into the social phase of development.   Enzyme levels were depleted in the social phase of development by removing tetracycline from the media upon forcing the *D. discoideum* into the social phase.

Results:

There was no visible effect on the development of *D. discoideum* when depleting DdiThg1 expression during the either social phase nor the vegetative phase of development. The cells aggregated and formed fruiting bodies at a comparable rate to the wild type strain.

Depleting DdiTLP3's expression during the social stage of development resulted in very slow growth, and an inability to form stalks with fruiting bodies. Conversely, knocking the enzyme down during the vegetative phase of development resulted in faster fruiting body formation, with larger aggregates than the wild type cells at 24 hours. However, the faster-growing DdiTLP3 knockdown aggregates looked similar to wild type stalks at 48 hours.

Depletion of DdiTLP4 expression during the vegetative and social phases of development had similar results to the DdiTLP3 knockdown strains, but with more extreme phenotypes. Depletion of TLP4 during the social phase resulted in a much slower formation of fruiting bodies, with several of the slugs being arrested in the early 'mound' stages of development. Knocking the enzyme down during vegetative phase of development resulted in faster formation of stalks and fruiting bodies, which were much larger than both the wild type and the TLP3 knockdown strains. Additionally, there was a higher proportion of aggregated *D. Discoideum* that exhibited abnormal morphology. Among these abnormalities were aggregated *D. discoideum* that failed to form a fruiting body and only consisted of long, thick stalks; aggregated *D. discoideum* that formed two stalks connected to one fruiting body; and aggregated *D. discoideum* forming fruiting bodies that were significantly longer and thicker overall. A table summarizing the developmental phenotypes of the different knockdown strains can be found in Table 1**,** and pictures of the different knockdown strain phenotypes can be found in figure 7**.**

CHAPTER 2: *In Vitro* Characterization of DdiTLP4

Measuring N$_{-1}$ Addition by DdiTLP4 to Minimal tRNA^{His} Substrates

Single-turnover assays were performed in an effort to characterize the kinetic properties of

DdiTLP4 and the Thg1/TLP family as a whole. Characterizing the kinetic properties of DdiTLP4 on

different substrates can help illuminate properties unique to DdiTLP4 in terms of its binding

preferences and relative reaction rates. More importantly, characterizing DdiTLP4's kinetic

properties helps provide a greater overall understanding of what Thg1/TLP family members

prefer to act on, and by extension bind to, in a substrate. This knowledge could help to

determine identity elements that the enzyme family members use to recognize and interact with

their various RNA substrates, thus facilitating identification of new in vivo substrates for 3'-5'

polymerases. Single-turnover assays were performed on minimal RNA substrates to assess how

well TLP4 catalyzes nucleotide addition [Figure 8]. Both Watson-Crick and non Watson-Crick

nucleotide addition reactions were tested. Previously collected data from ScThg1 and other TLPs

from two different domains of life, bacterial *Bacillus thuringiensis* BtTLP and archaeal *M. smithii*

TLP (MsTLP), was used as a control.

Minimal RNA substrates were used in these single-turnover assays to help elucidate the

basic characteristics that the Thg1/TLP enzyme family recognizes as identity elements in a

substrate. Each of these four substrates differed from one another by one property, such as

length or sequence, so that any observed difference in relative rate could be attributed to a

specific aspect of the substrates' identities. The minimal RNA substrates used were created were

designed to match a tRNA molecule's acceptor stem region, guided by the idea that RNA

molecules can be aminoacylated if they have structures that match the acceptor stem region of a

tRNA molecule [6]. Substrates of two varying lengths were created, in order to test what size the enzymes preferred. The substrates were designed to match either tRNA^{His}'s 12 base pair acceptor stem and T-loop region ("mini N$_{73}$ substrates"), which fold coaxially over each other as a mimicry of the structure of real tRNA$_{His}$ [figure 1] [figure 8], or to match only molecule's 7 base pair acceptor stem ("micro N$_{73}$ substrates"). Additionally, the substrates were designed with the discriminator nucleotide (N$_{73}$) in mind, as eukaryotic tRNA^{His} has A$_{73}$ while prokaryotic tRNA^{His} has C$_{73}$. For more details on each of the minimal RNA substrates and their sequences, see [figure 8].

The RNA substrates' 5' ends were labeled using [γ-^{32}P]GTP, thus generating minimal His substrates labeled at the terminal phosphate at their 5' ends. Nucleotide addition to the labeled substrates was tested using a nucleotidyl transfer assay, where the substrates were incubated at room temperature for 120 minutes with DdiTLP4, and 5uM of either GTP or UTP. As the incoming N$_{-1}$ attacked the RNA 5' end of the substrate, the labeled P was released as a radioactive pyrophosphate (^{32}P*P$_i$), which could then decay over time to form a radioactive inorganic phosphate (^{32}P*) [figure 9]. Aliquots were taken at individual timepoints, treated with EDTA to stop the reaction and with RNase A to digest the labeled substrate into smaller pieces. TCA was added to precipitate any enzyme or unreacted RNA that may interfere with resolving the final product on a TLC plate. The samples were then resolved by PEI cellulose-TLC.

Results:

The observed reaction rate for a given nucleotide concentration (k$_{obs}$) corresponds to the amount of nucleotide additions the enzyme catalyzes per minute at the given concentration of nucleotides. As can be seen in figure 10, DdiTLP4 displays a preference towards acting on the longer, 12 base pair mini stem loops as opposed to the shorter, 7 base micro pair stem loops.

The enzyme was incapable of catalyzing non-Watson Crick base paired additions, such as G-A$_{73}$ or U-G$_{73}$ [Table 3]. These characteristics line up with the other TLPs that have previously been studied, and help to create a clearer picture of the binding preferences of the TLPs in general [7] [8]. However, DdiTLP4 differs from BtTLP and MsTLP in that it prefers to add GTP across from C$_{73}$ as opposed to adding UTP across from A$_{73}$ and has extremely low activity on the micro A$_{73}$ substrates, where BtTLP and MsTLP seemed to prefer the A$_{73}$ substrates. DdiTLP4 also had a much faster reaction rate than ScThg1 and the other TLPs. Attempts to perform the reaction at NTP concentrations of 1mM, the concentration of NTP used when testing the other TLPs and Thg1, resulted in near saturation and over 70% of product formation within the first 30 seconds of the reaction. Various single turnover assays were performed with GTP and UTP concentrations between 0.1uM to 1 mM, and it was found that reactions involving DdiTLP4 and C$_{73}$ substrates only need 5uM of NTP to become saturated within the allotted timeframe, and to achieve similar k$_{obs}$ results to those of reactions involving Thg1 and the other TLPs at 200uM of NTP. Thus, DdiTLP4 is capable of comparable observed rate constants to the other TLPs at nucleotide concentrations 200 times smaller than those used to test the other enzymes' activities, and is a much faster enzyme overall. A table comparing the relative observed rates of DdiTLP4, ScThg1, BtTLP and MsTLP for the four minimal RNA substrates can be found in table 2.

CHAPTER 3: *In Vitro* Characterization of DdiTLP4

Structure-Function Assays of Various TLP/Thg1 Mutant Enzymes

Determining what specific features of DdiTLP4's structure give it its substrate specificity and allow it to edit its (hypothesized) non-tRNA substrates could provide valuable information about the mechanism Thg1 and the TLPs utilize to recognize and bind to their specific substrates.

Mutant DdiTLP4 enzymes, as well as mutant ScThg1 and BtTLP enzymes, were generated in an attempt to characterize the relationship between structure and function of DdiTLP4, while concurrently assessing ScThg1 and BtTLP.

Two complementary approaches were used to generate mutant enzymes. In one method, mutations were made based off of the structure of DdiTLP4.   The goal of this method was to create loss-of-function mutations in DdiTLP4 and BtTLP (which can catalyze the same reactions as DdiTLP4 *in vivo*) while creating simultaneous gain-of-function mutations in ScThg1, in terms of the enzymes' substrate specificity and their ability to catalyze nucleotide addition to 5srRNA or ncRNA. Target residues of DdiTLP4 were replaced with the corresponding residues in ScThg1. The target residues in BtTLP were also replaced with the corresponding residues in ScThg1, and vice versa. For some mutants AcaTLP2 was used in place of DdiTLP4, as AcaTLP2 performed the same reactions as DdiTLP4 *in vitro* and was easier to purify in higher amounts.

Potential target residues were first identified in a sequence alignment between DdiTLP4 and two enzymes that can catalyze nucleotide addition to 5s rRNA and ncRNA, BtTLP and *A. Castellani* TLP2 (AcaTLP2), and 11 other TLPs that could not catalyze said reaction [figure 11]. Regions or amino acids that were conserved among TLPs who can add nucleotides to 5srRNA/ncRNA were identified as candidate regions. To determine which candidate region to test first, the candidates were located on a crystal structure of BtTLP and prioritized based off of how close or far they were from an active site. The structure of BtTLP is functionally similar to that of Thg1, and it was assumed that DdiTLP4 would have similar structure as well [figure 13]. The target residues were then mutated using a Phusion PCR method where the mutation was incorporated into the PCR oligo template. The enzymes, which all were designed to contain His

tags, were then purified using immobilized metal-ion affinity chromatography (IMAC) on a cobalt resin, and their activities were tested using an *in vitro* enzyme activity assay called the phosphatase protection assay. The substrate, a 5s rRNA truncated before its terminal 5' G with a non-phosphorylated U at its 5' end (5s rRNA ΔG), was treated with PNK in the presence of [γ-^{32}P*]ATP to generate monophosphorylated UTP-rRNA, labeled at its phosphate. The substrate was then incubated at room temperature for 120 minutes with the enzyme being tested, ATP, and GTP. The incoming G_{-1} formed a phosphodiester bond with the labeled 5'- UTP, thus generating G^{32}p*U-rRNA. The substrate was then quenched with EDTA and treated with RNase A in the manner described in the nucleotidyl transfer assay, thus generating a G^{32}p*U fragment [Figure 16]. The reaction was then treated with CIP to remove the radioactive phosphate from any unreacted ^{32}p*U substrate, and resolved on a TLC silica plate [figure 17].

Two candidate regions were tested. One region consisted of histidine 145, a residue conserved between 5s/ncRNA editing enzymes that located in close proximity to the three carboxyl-group nucleotide transfer active site [figure 13A]. ScThg1 has a tyrosine at the corresponding location, so the mutations made were H145Y in AcaTLP2 (which could be purified in higher quantities than DdiTLP4) and Y129H in ScThg1. The second candidate region consists of two residues, TL, that are found within a variable loop connecting alpha helix D to beta loop 3 [figure 13B]. The variable loop is relatively close to the enzyme's ATP binding site. The loop regions were swapped between DdiTLP4, BtTLP, and ScThg1 in the manner described above.

The second method used revolved around generating different combinations of double mutant Thg1 enzymes based off of previously known mutations that affected Thg1's substrate specificity, relative reaction rates, or reaction efficiency. Previous research has shown that three

point mutations to ScThg1 (D68A, E179A, D153A) lead to expanded substrate recognition in the ScThg1. Creating a triple mutant of the three regions, as well as a D68A and E179A double mutant, have also been shown to increase ScThg1's substrate recognition and also its activity levels. However, the effects of combining the D153A and E179A, as well as the D68A and D153A mutations in ScThg1 were unknown. Using the D153A ScThg1 plasmid as a template, a second mutation was incorporated into the enzyme's sequence using the Phusion PCR method described above. The resultant double-mutant enzymes were then tested in the same way as the TLP/Thg1 single mutants. See table 3 for a list of all the different mutations that were tested.

Results

There were several problems in the PCR used to generate the H154Y/Y1129H active site mutants, as well as in the sequencing step used to validate the mutations, and the mutation could therefore not be generated for testing within the necessary timeframe. The variable loop mutants could be cloned and purified, but mutating the variable loop in DdiTLP4 and ScThg1 disrupted the enzymes' structures, causing them to be insoluble. When the different fractions generated during the protein purification process were run using gel electrophoresis on an SDS-PAGE gel, there were very low amounts of enzyme found in the soluble fraction of the protein purification process and even lower amounts of enzyme found in the purified final product [figure 14]. Mutants for which significant quantities of soluble purified protein could be obtained were also tested alongside the BtTLP mutant, but none of the mutants showed any activity in the phosphatase protection assay [figure 18].

In the same vein, SDS-PAGE gel analysis reveals that the D153A + D68A double mutant could not be expressed by its host enzyme, and therefore could not be tested [figure 14]. The

D153A + E179A double mutant, however, was successfully obtained and expressed in the E. coli

cell. The D153A + E179A double mutant exhibited both an expanded substrate specificity and an

increased level of overall activity, as enzyme concentrations of the double mutant lower than the

concentration used to test BtTLP were shown to be capable of adding G_{-1} to 5s rRNA, and

produced a more intense product spot than the BtTLP enzyme.

Discussion/Future Directions

Two conclusions can be drawn from the *in vivo D. discoideum* knockdown experiments.

First, DdiThg1 did not display any developmental phenotypes when its levels are reduced, and

therefore does not play a role in the aggregation stage of the organism's development. Second,

DdiTLP4 and/or its substrates seem to have different functions during the two different phases of

D. discoideum development. Third, DdiTLP3 and/or its substrates seem to play a role in the

vegetative state of *D. discoideum* development. DdiTLP3 aggregates grew faster when enzyme

levels were reduced in the vegetative stage and could not properly aggregate when enzyme

levels were reduced in the social phase. However, further testing is necessary to determine

whether or not DdiTLP3 plays a different role during both stages of development. This is because

by 48 hours the faster-growing aggregates had stopped growing and were morphologically

similar to the wild-type strain, and because the experiment was only carried out one time. The

faster growth rate could have been a coincidence, and may be due to individual variation

between cells. Therefore, a future direction would be to repeat the TLP depletion tests more

times and attempt to quantify the results. The similar phenotype caused by depleting levels of

DdiTLP3 and DdiTLP4 during the social phase of development are most likely driven by different

processes as the two enzymes localize to different parts of the cell (the mitochondria and the

cytoplasm, respectively) and edit different substrates [4]. Further investigation of the knockdown strains in general is required to understand the cause of these interesting mutant phenotypes. For example, it would be beneficial to perform RNAseq on the *D. discoideum*'s RNA in the vegetative and social knockdown strains for both enzymes, and compare the 5' ends of the different strains' RNAs to see what is and is not being edited.

Two conclusions can be drawn from the single turnover kinetic assays. First, the data obtained from DdiTLP4 nucleotide addition on the minimal RNA substrates reinforces the previously obtained conclusions that TLPs prefer to add nucleotides to longer RNA substrates while Thg1 prefers the shorter substrates, and that TLPs are less proficient at of carrying out non-Watson Crick nucleotide addition. Second, DdiTLP4 is a much faster enzyme than ScThg1 and the other tested TLPs, obtaining similar relative rate values to the aforementioned enzymes at nucleotide concentrations 200-fold smaller than those used to test ScThg1 and the other TLPs. This faster reaction rate expands the known limits of how quickly a TLP or Thg1 is capable of catalyzing nucleotide additions, and may indicate that DdiTLP4 has lower nucleotide k_d than Thg1 or BtTLP MsTLP. Future directions should involve testing DdiTLP4's performance in single-turnover kinetics assays on the minimal RNA substrates using various concentrations of NTP. Third, DdiTLP4 prefers to catalyze C-G base paired reactions over U-A base paired reactions.

Two conclusions can be drawn from the biochemical characterization of DdiTLP4's structure. First, the variable loop region present in all three enzymes is not involved in substrate recognition and cannot be too dramatically changed without rendering the mutant protein insoluble. While the BtTLP loop mutant was soluble and technically underwent a loss-of-function phenotype, the fact that the other two mutants' structures were so disrupted leads us to believe

that this loss-of-function is related to a structural disruption in BtTLP as opposed to an actual

alteration in the regions that regulate substrate specificity. Second, the combined roles of the

residues D153 and E179 of ScThg1 play roles in substrate recognition and activity levels of the

enzyme. No concrete conclusions can be made about the other double mutant, as it could not be

expressed. Interestingly, none of the three mutated residues in ScThg1 (D68, E179, D153) are

notably close to either of the enzyme's active sites, indicating that other regions of the enzyme

could play a role in substrate recognition and reactivity. Future directions for this project include

generating and characterizing the histidine/tyrosine active site mutants, the D68A + D153A

double mutant, as well as more mutant TLPs based off of the target residue list.

Materials and Methods

Growing *D. Discoideum*

On the first day, frozen *D. discoideum* sample strains were partially thawed 40 seconds in a 37 ^{o}C

water bath and transferred into 15 mL conical vials containing 10 mL of HL5 buffer. They were

then incubated 30 min at 21.9^{o}C for recovery, and spun down at 500g for 3 min and suspended

in fresh 10 mL of fresh HL5 to wash off the DMSO in their save media. 24 uL of 12.5 mg/mL

tetracycline was added to each tube at a final concentration of 30ug/uL to prevent the RNAi

from knocking down its respective genes. The *D. discoideum* strains were then plated on

bacterial culture dishes and incubated O/N at 21.9^{o}C. On day two, various antibiotics were added

to the different 10 mL plates. The samples were then left to incubate again overnight at 21.8^{o}C.

G418 and blasticidin S were added to all four knockdown strains at final concentrations of 20

ug/mL and 10 ug/mL, respectively. On day three, each confluent 10 mL bacterial plate of *D.*

discoideum was split into two 50 mL bacterial plates containing 40 mL of HL5 medium, 30 ug/mL

of tetracycline, 20 ug/mL of G418, and 10 ug/mL Blasticidin S, and incubated for 48 hours at 21.8

21.8⁰C. Upon reaching confluence, each sample was transferred from bacterial dishes to two 50

mL conical tubes, spun down for 2 min at 500 g, and then resuspended and counted. The

samples were then washed in 50 mL developmental buffer 3x, resuspended in 300 uL DB, and

plated on KK2 plates at a density of 7.4E7 cells per plate. The samples were incubated at 22⁰C,

and pictures were taken of the plates 24 and 48 hours into their development. The recipes for

developmental buffer, HL5 media, and KK2 plates can be found in [9].

Preparation of the mini and micro His C73 DNA substrates

The mini and micro C73 substrates' DNA were digested from their respective plasmids at 60°C

using a BstNI restriction endonuclease (fc 1mM), in the presence of 1mM 10x NEB buffer 3.1.

They were then purified with PCA and tris pH 8 at a ratio of 1:2:1 DNA:PCA:tris to remove organic

materials such as the enzyme, and the DNA was then precipitated in the presence of glycerol and

100% ethanol.

Preparation of the mini and micro His A73 DNA substrates

Plasmids containing the substrate were purifying using a Qiagen MaxiPrep kit, and then digested

using restriction endonucleases. while the A73 substrates were digested from their plasmid DNA

at 37°C in the presence of 10x NEB buffer #3.1 (fc 1 uM) for one hour using NsiI restriction

endonuclease (fc 1mm). They were then PCA purified and ethanol precipitated. The A73

substrates were then treated with T4 DNA polymerase (1U) in the presence of 1mM 10x NEB

buffer #3 and 100mM dNTPs to remove overhanging fragments left behind by NsiI, thus creating

blunt ends more suited for a run-off transcription.

***In Vitro* transcription of radioactive 5'- [γ-³²P] GTP mini and micro RNA substrates**

All four substrates were cloned into their respective E. Coli plasmids under the control of a T7 promoter, which is recognized by the promoter-specific T7 RNA polymerase. The DNA substrates were incubated at 37°C for two hours in the presence of T7 RNA polymerase, NTPs, and radioactive [γ-^{32}P]GTP for two hours. DnaseI was then added to the reactions, which incubated for another 30 minutes as the DNA templates were digested into individual NTPs. The RNA was then purified by gel electrophoresis on a 12% polyacrylamide 7 M urea gel run at 15 W for 40 minutes, and the RNA bands were visualized and cut out from the gel under a UV light. The RNA was then eluted from the gel using an RNA elution buffer, PCA purified, ethanol precipitated, and resuspended in 20uL tris pH 7.5.

Single-turnover kinetics assay

30 uL reactions were carried out in 25 mM HEPES pH 7.5, 10mM MgCl$_2$, 3mM DTT, 125 mM NaCl, 0.2 mg/mL BSA, 15 mM of the enzyme being tested, and 200 cpm/uL of the labeled RNA substrate. The reaction mixture incubated at room temperature (~21°C) for 120 minutes, with 2 uL aliquots taken at 0.5, 1, 2, 5, 10, 15, 30, 60, 90, and 120 minutes. Each aliquot was mixed with 1 uL of a 1:1 EDTA:RNase A mixture. The EDTA was used to sequester the metal ions in the enzyme's active site and stop the reaction, and the RNase A was used to ensure that any unreacted 5'[γ-^{32}P]GTP was separated from the rest of the RNA . The aliquots were incubated for 10 minutes at 50°C to activate the RNase A. The reactions were then incubated at 4°C for five minutes with trichloroacetic acid to precipitate the enzyme, and spun at 4°C for five minutes to remove any unreacted product. The resultant supernatant was spotted onto PEI cellulose-TLC plates, which were then soaked in methanol to activate them, and run on an 80:20 KPO$_4$:MeOH solvent system. The reaction was then visualized using a Typhoon phosphoimager, and percent

product formation and the observed reaction rate was measured and analyzed using ImageQuant and Kelidograph.

Purifying DdiTLP4, BtTLP, ScThg1 wild type and mutant enzymes

BtTLP and ScThg1 mutants were transformed into BL21(DE3)/PLys E. Coli cells for overexpression, while DdiTLP4 mutants were transformed into Rosetta/PLys E. Coli. The cells were cultured, and the His tagged proteins were purified using metal-ion affinity chromatography (IMAC) on cobalt beads. The proteins were dialyzed with 50% glycerol, quantified using a NanoDrop, and stored at -20°C. Purity and presence of the protein was tested using gel electrophoresis on an SDS PAGE gel

SDS Page gel

Each fraction collected during the gel purification process was diluted and run with dye (dye:BME 1:10) on a 12% SDS-PAGE gel at 180V for 50 minutes and stained for 30 minutes. The gel was made up of a resolving gel and a stacking gel. The resolving gel consisted of 40% Acrylamide (29:1), 1.5M Tris pH 8.8, 10% SDS, 10% APS, TEMED, and ddH$_2$O. The stacking gel was made with 40% Acrylamide (29:1), 1.0 M Tris pH 6.8, 10% SDS, 10% APS, TEMED, and ddH20. The gel was then stained for 30 minutes, and destained overnight in 1:1 destain solution:ddH$_2$O. Resultant bands were compared to the expected protein size to ensure that the correct protein, and only the correct protein, was obtained.

Cloning mutant enzymes using Phusion PCR

Primers were generated that matched the sequence surrounding the area to be mutated in the enzyme, and differed from the template DNA with respect to the areas meant to be mutated. The primers were phosphorylated using T4 PNK. PCR conditions were 1x of 5x Phusion buffer, 50

ng template DNA, 0.5 uM of the forward primer and 0.5 uM of the reverse primer, 1 unit of 2

U/uL Phusion HS polymerase, and 200uM dNTP. The first five cycles consisted of the

denaturation step, which occurred for 1.5 minutes at a temperature of 98°C, the annealing step

was 30 seconds at 55°C, and the short extension step was 3 minutes long at 72°C. The next 25

cycles consisted of 10 seconds at 98°C followed by 30 seconds at 65°C, and 3 minutes at 72°C.

The final cycle consisted of 10 minutes at 72°C. The resultant PCR product was ligated together

to form a whole plasmid, transformed into XL1- Blue E. Coli cells, and grown on LB/Amp plates at

37°C. Colonies were sent to GeneWiz for sequencing, and the DNA of positive clones was

isolated using a Qiagen MiniPrep kit.

In vitro Hammerhead Transcription of truncated 5s rRNA Substrates

Plasmids containing the substrate were purifying using a Qiagen MaxiPrep kit, and digested with

the restriction endonuclease Hpy991, and then treated with T4 DNA polymerase to ensure the

DNA ends were blunted for a runoff transcription. The DNA was then mixed with 80 mM trisCl

pH 7.5, 1mM spermidine, 30 mM MgCl2, 5 mM DTT, 0.12 mM BSA, 4 mM of ATP, CTP, UTP and

GTP, and ddH$_2$O. The reaction mixture was incubated at 37°C for 2 hours with T7 DNA

polymerase until the solution was turbid. The RNA then underwent thermo cycling for 10 cycles

of 3 minutes 60°C + 3 minutes 25°C for the hammerhead ribozyme in the RNA to undergo self-

cleavage, thus cleaving itself and the terminal 5' G. The RNA was then treated with DnaseI to

digest the DNA template, and purified on a 10% PA 4 M urea gel at 15W for 45 minutes in the

same manner as the minimal RNA substrates. The RNA was then quantified using a NanoDrop.

Labeling ^{32}p*UTP-5s rRNA substrates

40 pmol of substrate was reacted with [γ-^{32}P]ATP, 10x T4 PNK buffer, T4 PNK, and ddH2O at 37°C for 60 minutes. The T4 PNK was then heat-killed at 72°C for 10 minutes, and the reaction was spun through a P6 gel column to remove any unreacted [γ-^{32}P]AT. The RNA was then mixed with an equal volume of 2X RNA dye, and purified on a 10% PA 4 M urea gel at 15 W for 45 minutes in the same manner as the

Phosphatase Protection Thg1/TLP Activity Assay

5000 cpm/reaction of labeled 5s rRNA was incubated at room temperature for 120 minutes with 25 mM HEPES pH 7.5, 10 mM MgCl2, 3 mM DTT, 125 mM NaCl, 0.2 mg/mL BSA, 0.1 mM ATP, 1 mM GTP, and ddH$_2$O with either 20 or 4 uM of enzyme (4 uM enzyme for the double mutant, and 20 uM of enzyme for all of the other mutant and wild type enzymes). After 2 hours, 1 uL of a 1:1 EDTA:RNase A mixture was added to each reaction, and the reactions were incubated at 50°C for 10 minutes. The reactions were then treated with 1 uL of a 1:19:20 CIAP:CIP dilution buffer:dephosphorylation buffer mixture, and incubated at 37°C for 1 hour. 2 uL of each reaction was then spotted onto a silica TLC plate and run overnight in a 55:35:10 n-propanol:NH$_4$OH:H$_2$O solvent system. The reactions were visualized using a Typhoon phosphoimager and analyzed using the ImageQuant program.
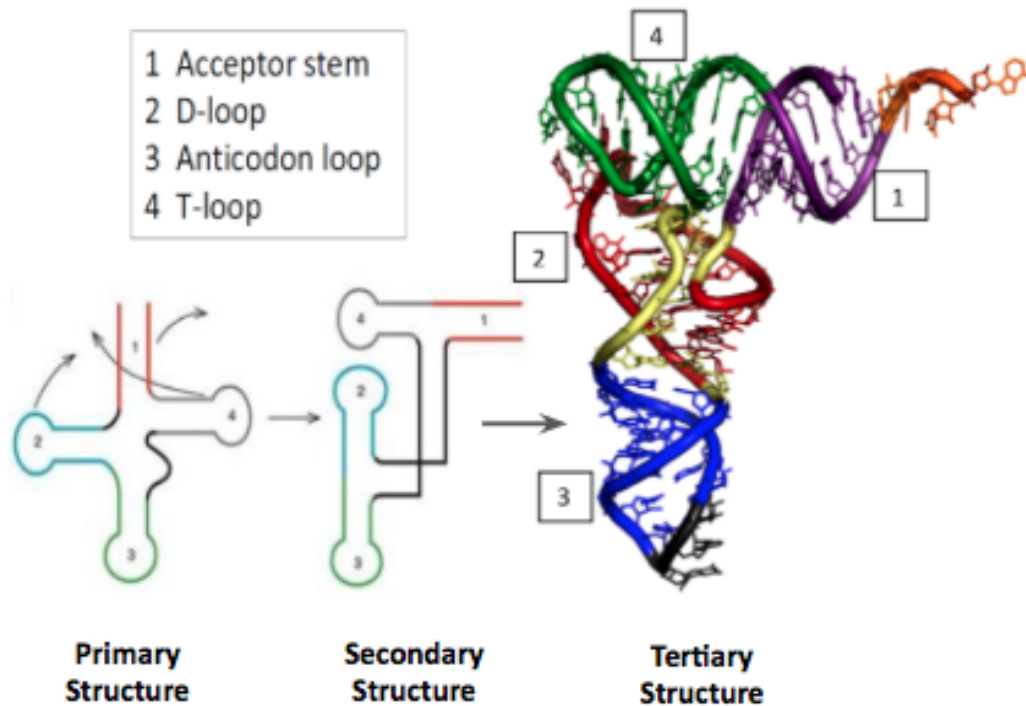
Figure 1: The Structure of tRNA. The acceptor stem of tRNA binds to and interacts with amino acids, the D-loop acts as a recognition site for aa-RS enzymes [10], the anticodon loop interacts with mRNA and helps determine which amino acid is incorporated into the growing polypeptide chain, and the T-loop acts as a recognition site for ribosome binding. In the tRNA's tertiary structure, the acceptor stem and the T-loop fold so that they are coaxially on top of one another. [Figure made by Lauren Duff] [7]
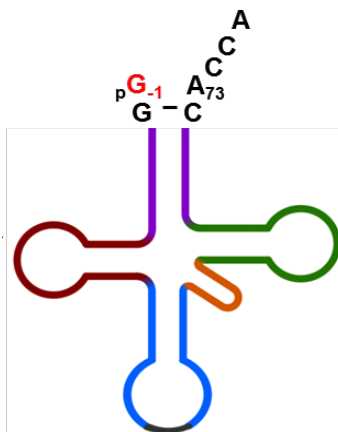

Figure 2: Mature tRNA^{His}. Mature tRNA^{His} has a G_{-1} on the 5' end of its acceptor stem, in a non Watson-Crick base pair with the opposite A_{73} nucleotide. [Figure courtesy of Dr. Yicheng Long]
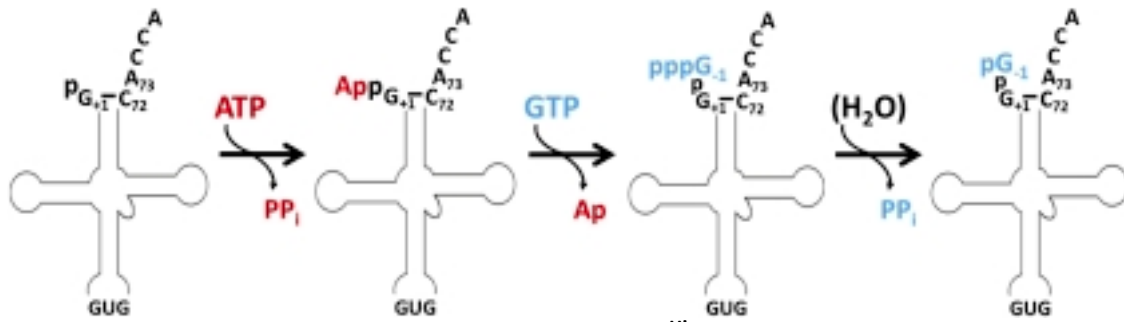
Figure 3: Three-step mechanism of G$_{-1}$ addition to tRNA^{His}, as catalyzed by Thg1. Nucleotide addition involves an adenylation step to activate the tRNA, a nucleotide transfer step, and a phosphatase step. [11]

A

B
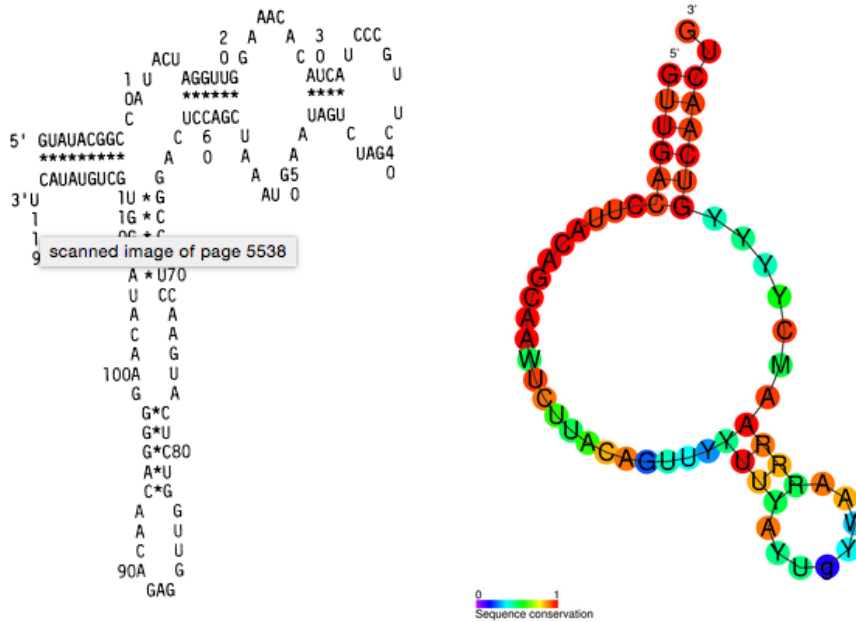
Figure 4: A. *D. discoideum* 5s rRNA. The 5s rRNA is approximately 120 nucleotides long and makes up the 5s subunit of the ribosome [11]. **B. *D. discoideum* class I ncRNA.** Class I noncoding RNA is 57 nucleotides long. Its function is unknown, but its expression is down regulated during development [3]. In this figure, warmer colors (red, yellow) correspond to nucleotides that are conserved among different types of class I ncRNA molecules, while cooler colors (purples and blues) are less conserved across the classification and more unique to the specific ncRNA.
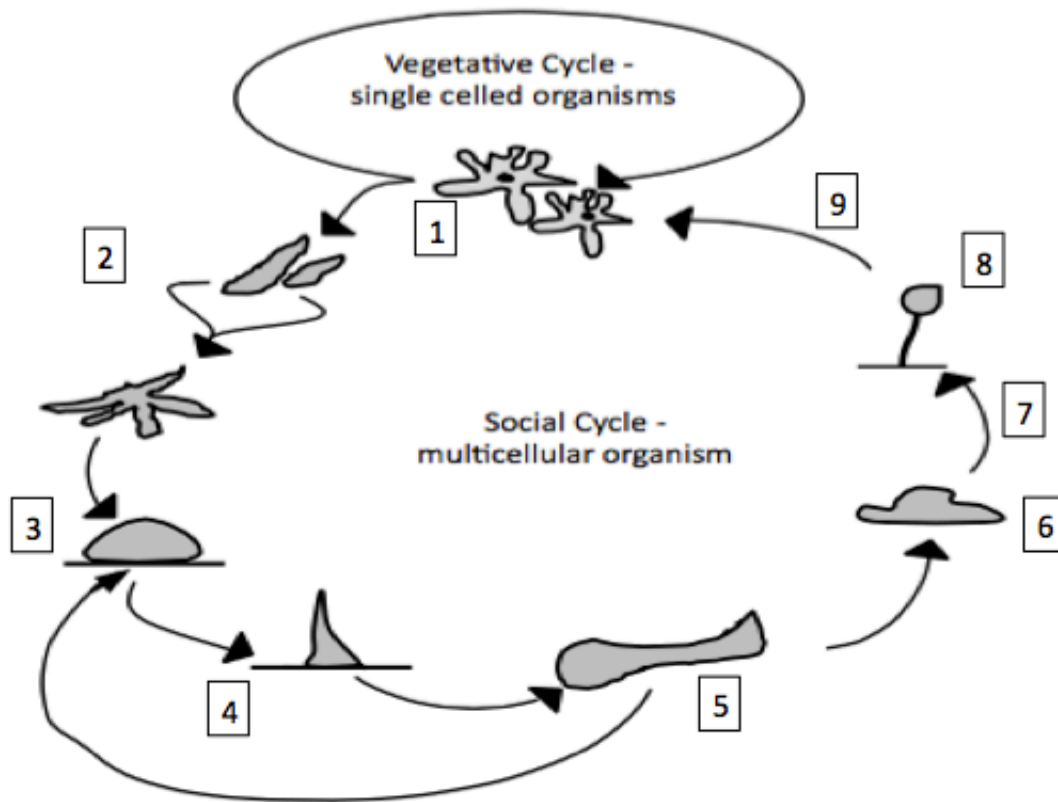
Figure 5: *D. discoideum* life cycle. *D. discoideum* begins its life cycle in the vegetative phase, as single celled amoeba. Upon starvation they aggregate together to seek food and, if food cannot be found, reproduce. Counterclockwise from the top, the stages of *D. discoideum* social cycle are: **1.** The organism exists as a single-celled amoeba in the vegetative stage. **2.** Amoeba aggregate towards each other in a process called streaming. **3.** The aggregated cells form a mound. **4.** The mound changes form into a structure called the "first finger". **5.** The finger differentiates into a slug, which moves around to look for food. The slug can either revert back to a mound, or continue along the developmental cycle. **6.** The slug begins to differentiate further, and it forms structure called a Mexican hat. **7.** The cells enter the culmination phase, during which the anterior cells of the slug die and push upwards to form a slug, while the posterior cells of the slug differentiate into a fruiting body containing dehydrated *D. discoideum* spores. **8.** The *D. discoideum* now exists as a multicellular organism with a stalk and a fruiting body. **9.** The spores contained in the fruiting body (generated through both sexual and sexual reproduction) are released in the presence of a food source, and the spores hatch as single celled amoeba in the vegetative state. [9]
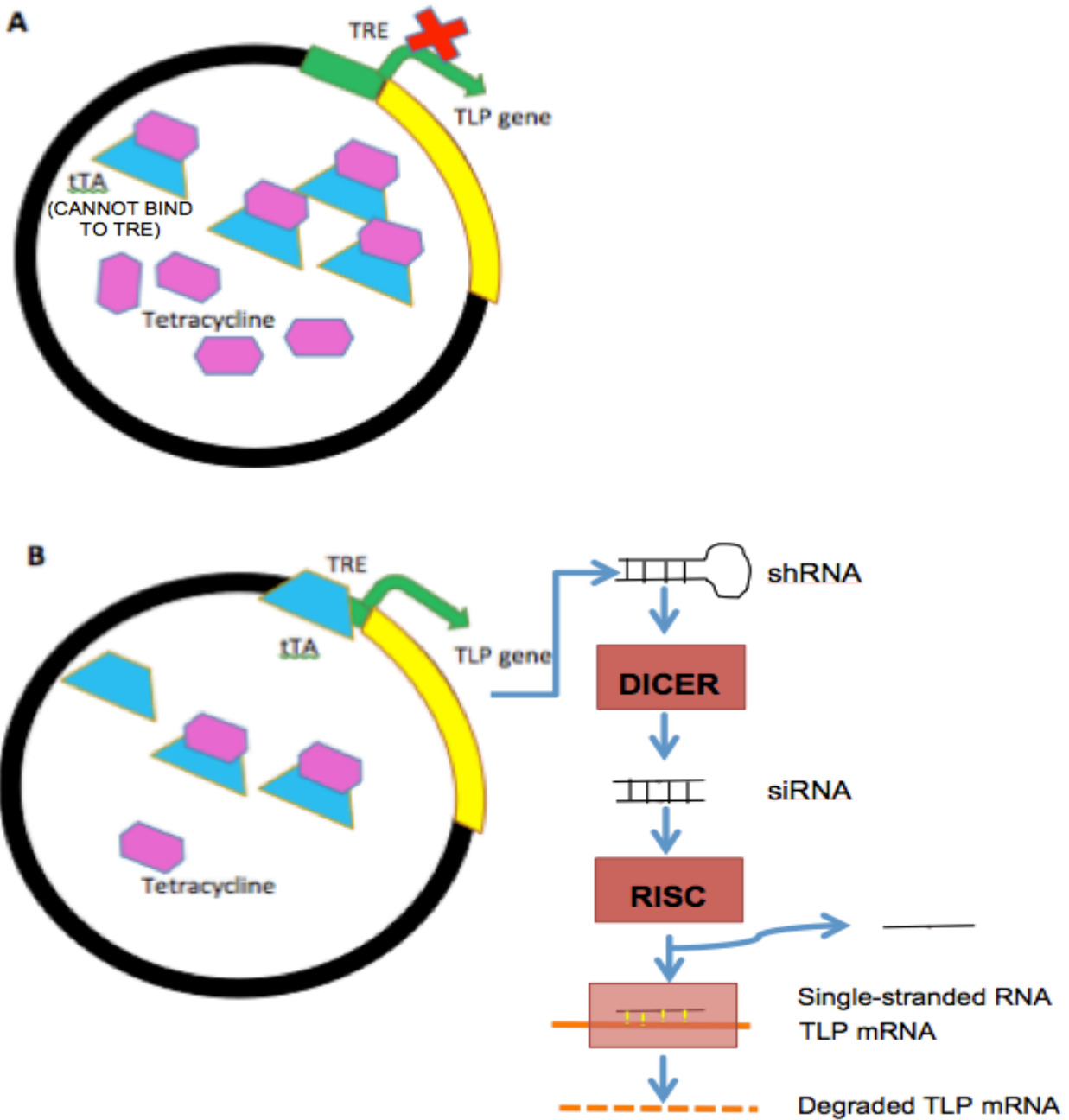
Figure 6: The Tet-OFF knockdown system and RNAi knockdown. A. The tTA preferentially binds to tetracycline over the TRE promoter region and the gene is not transcribed. **B.** In the absence of antibiotics, the tTA changes conformation and binds to the TRE promoter region. Transcription is activated. A short hairpin RNA (shRNA) transcript that is complementary to the message of the gene of interest is transcribed, and processed in the organism's endogenous DICER complex into double-stranded small interfering RNA (siRNA). The double-stranded siRNA is then fed into the RISC complex, where the strand that is not complementary to the gene of interest is removed and the RISC/complementary RNA complex binds to and silences the message of the gene of interest
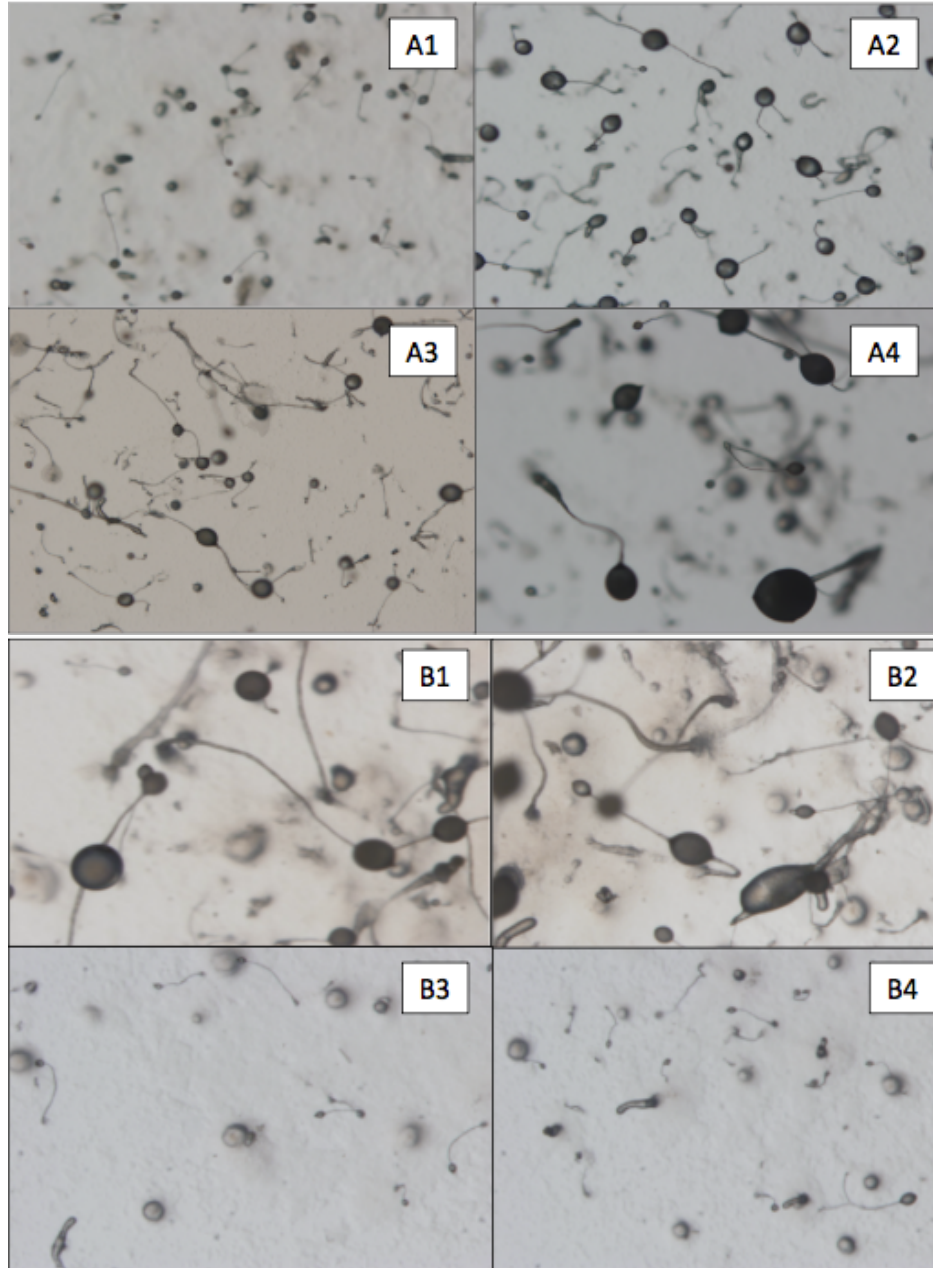
Figure 7: Developmental phenotypes of *D. discoideum* TLP knockdown strains. A. TLP knockdown in the vegetative phase, 48 hours after plating. **A1.** Wild type control cells. **A2.** DdiThg1knockout strain. There does not appear to be any effect. **A3.** DdiTLP3 knockout strain. Growth was faster than the wild type strain when checked at 24 hours (image not shown), but seemed to level out by 48 hours. **A4.** DdiTLP4 knockout strain **B.** TLP knockdown in social phase, 48 hours after plating. **BI.** Wild type control cells **B2.** DdiThg1 knockout strain. There does not appear to be much effect. **B3.** DdiTLP3 knockout strain. Aggregation is very slow, may be arrested in the mound stage. **B4.** DdiTLP4 knockout strain. Aggregation is very slow, may be arrested in the mound state.
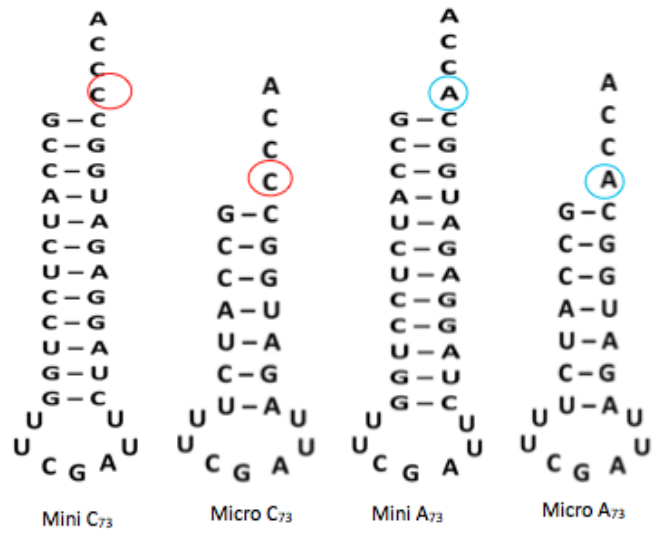
Figure 8: Minimal RNA substrates. Minimal RNA substrates were designed to mimic of *S. cerevisiae* tRNA[His]'s acceptor stem T-loop region [7]
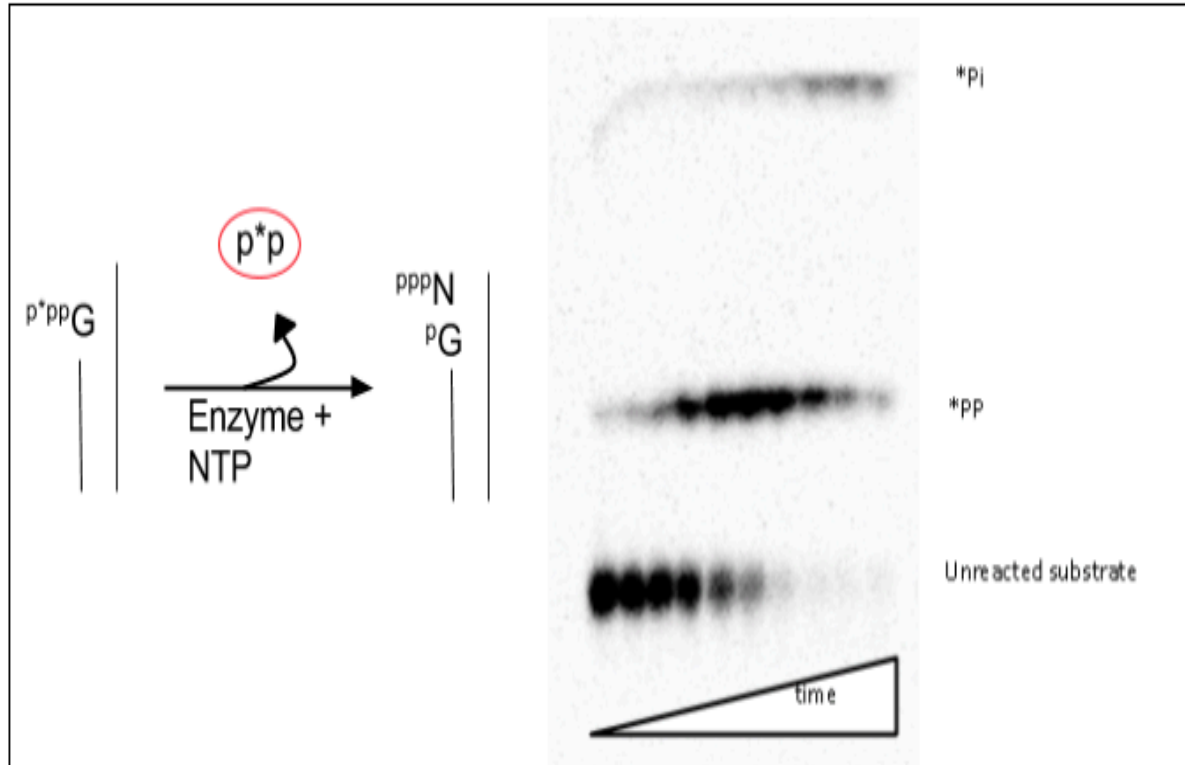
Figure 9: Nucleotidyl transfer TLP activity assay. Single-turnover kinetic assays were performed by DdiTLP4 on the mini A_{73} RNA substrate in the presence of UTP. If the reaction occurs, the incoming nucleotide will displace the two terminal phosphates as a radioactive pyrophosphate molecule. With time, the pyrophosphate will hydrolyze into two inorganic phosphate molecules, one radioactive and one not.
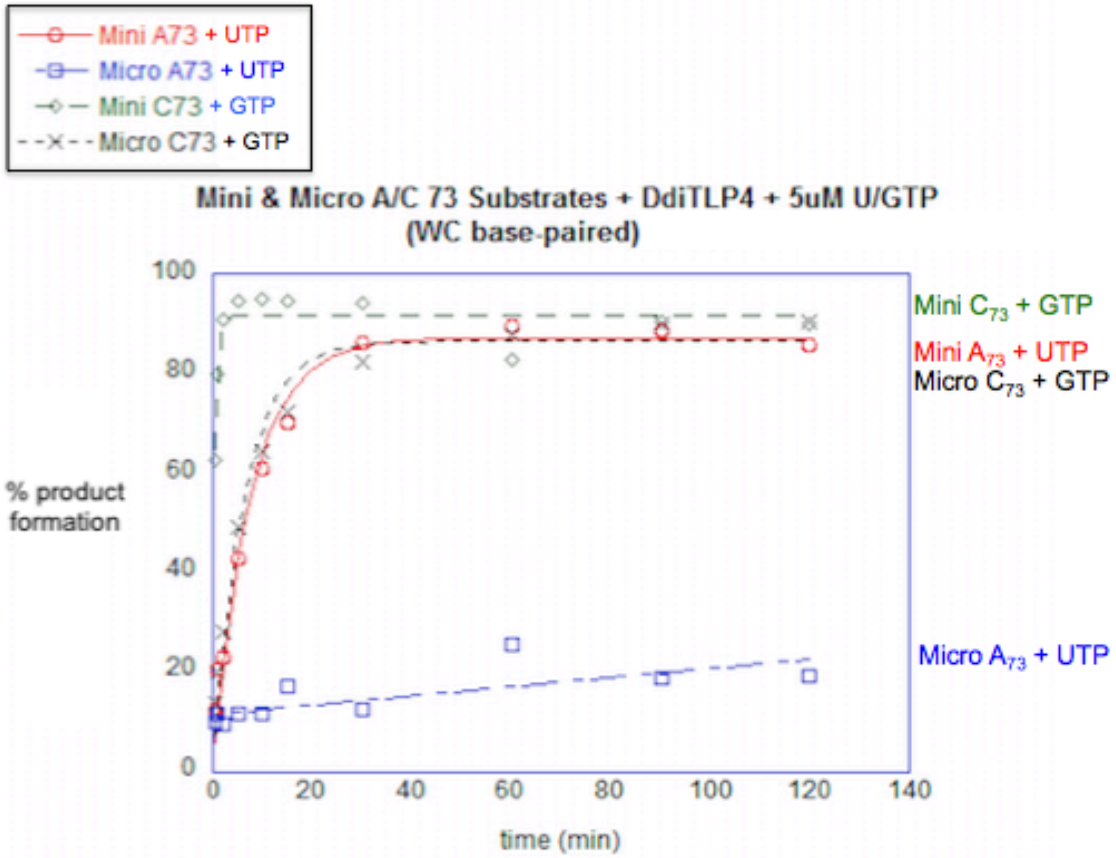
Figure 10: Nucleotide addition to minimal RNA substrates by DdiTLP4. DdiTLP4 exhibits faster reaction rates when acting on the longer mini-RNA substrates than the shorter micro-RNA substrates. It also exhibits faster reaction rates when acting on C_{73} substrates over A_{73} substrates.
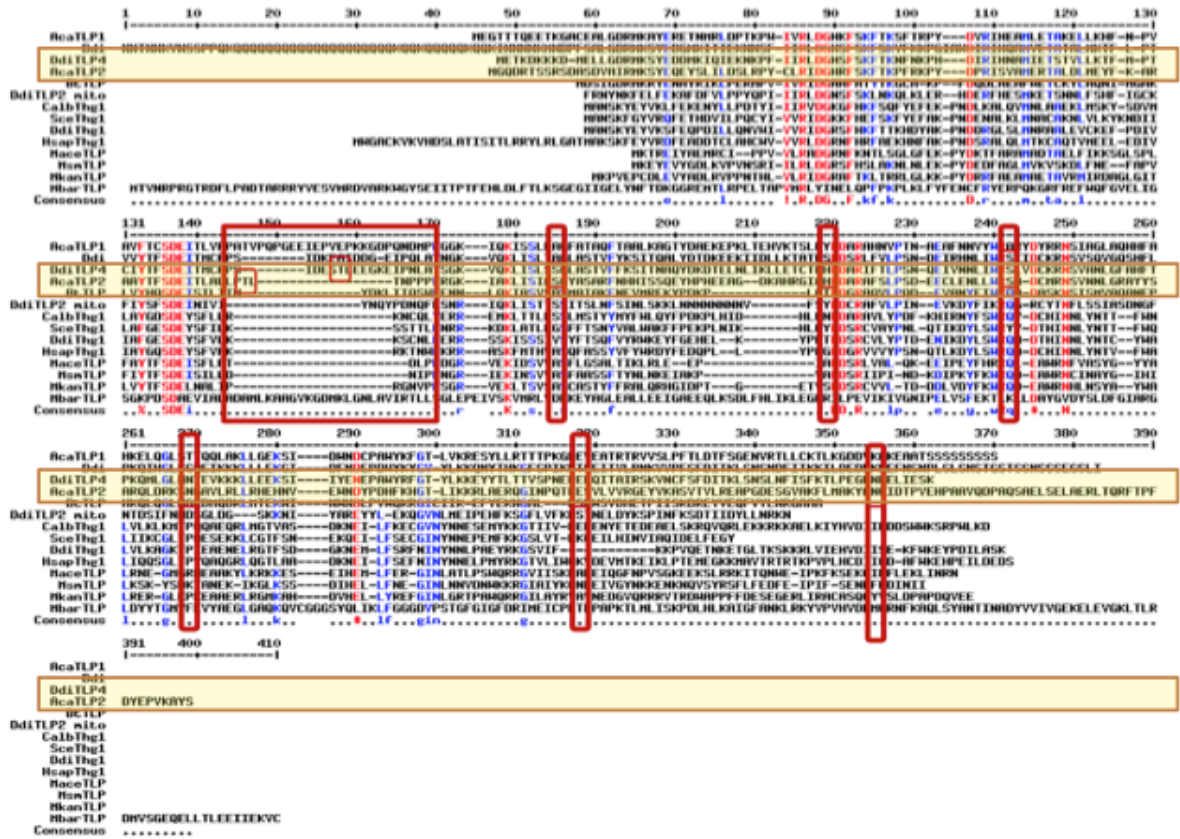
Figure 11: Sequence alignment to find potential ncRNA activity-granting residues. A sequence alignment was performed between DdiTLP4 and 13 other TLPs and ScThg1. The yellow-boxed sequences correspond to AcaTLP2 and DdiTLP4, which can add nucleotides to non coding RNA substrates *in vitro*. Residues conserved only among proteins that were known to act on ncRNA/5s rRNA (found in the red-boxed regions) were identified as potential candidates for mutagenesis.
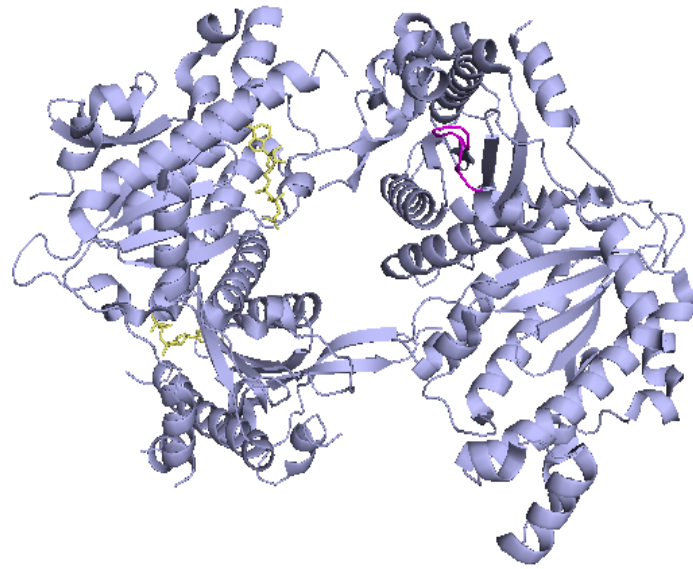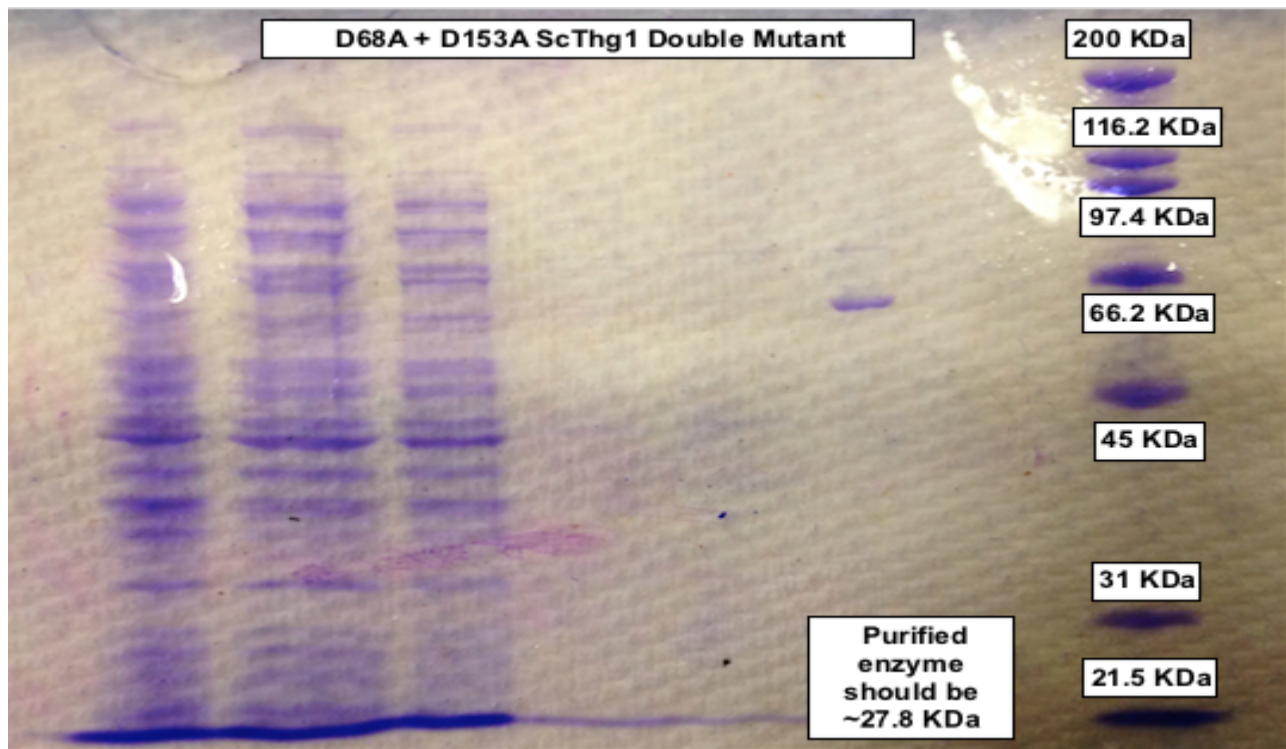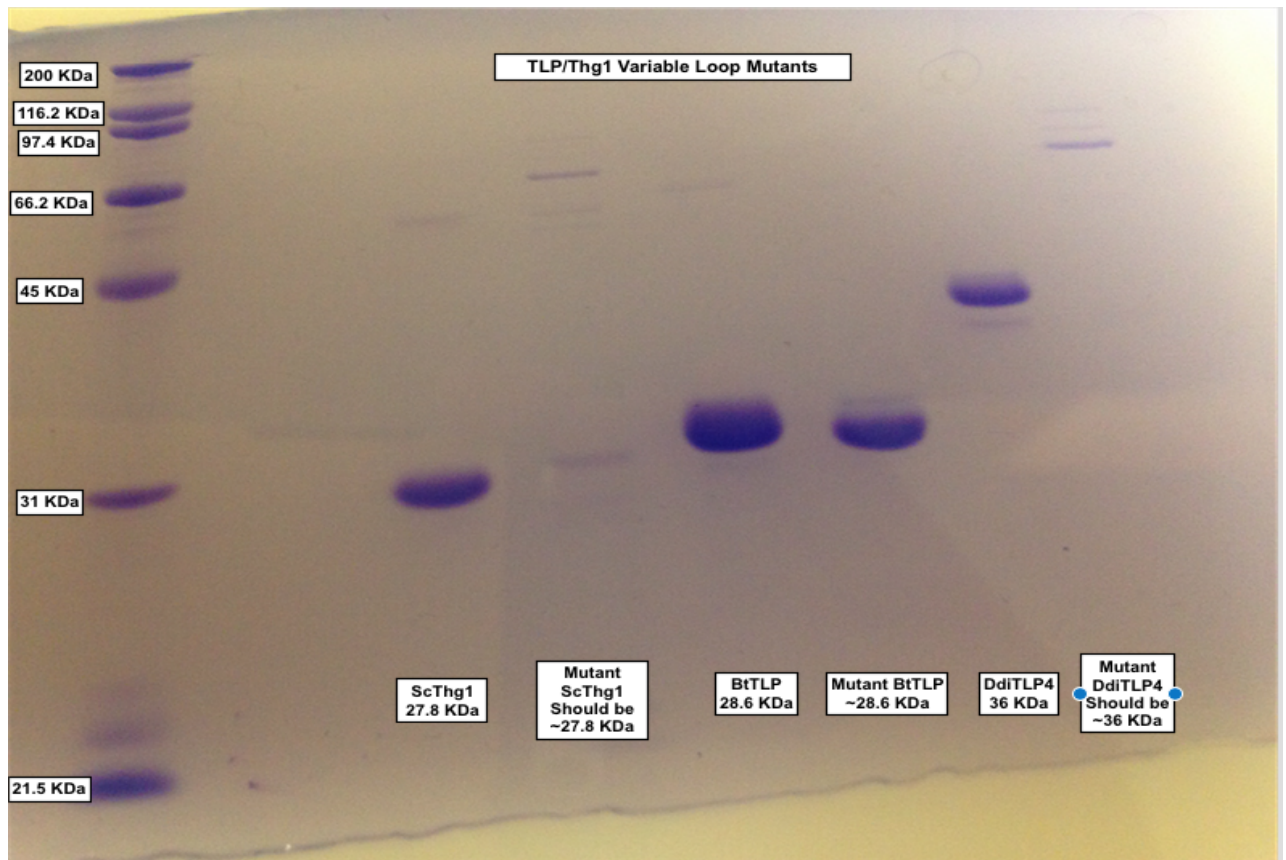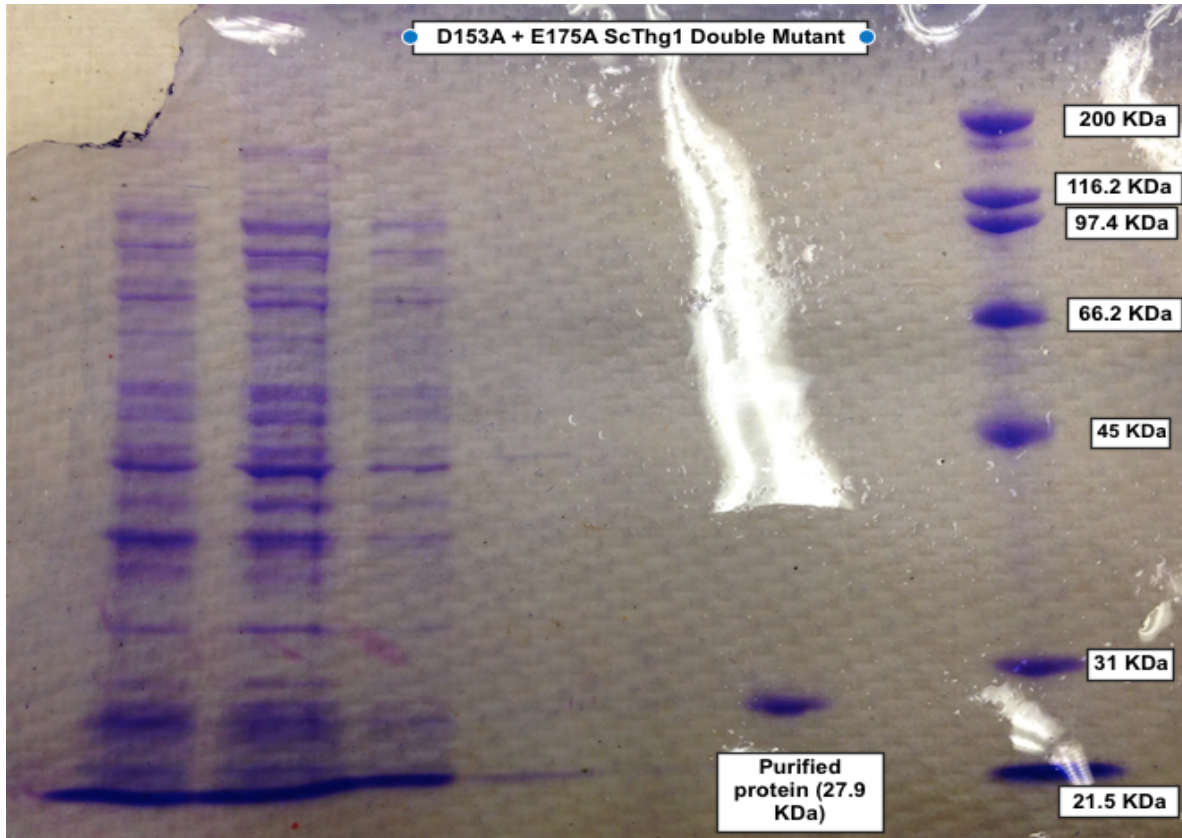
Figure 12: Crystal Structure of BtTLP. Enzymes in the Thg1/TLP family are made up of a dimer of dimers. They have ATP active sites and nucleotidyl transfer active sites, and coordinate two metal Mg^{2+} ions to maintain a metal ion catalysis mechanism. [12]
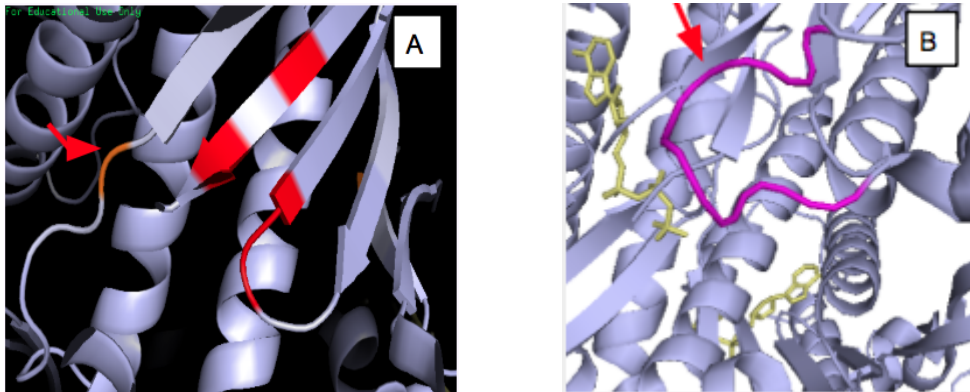

Figure 13: Mutated regions in DdiTLP4, BtTLP, and ScThg1. The two regions consist of **A.** a conserved histidine (orange) near the nucleotidyl transfer active site (red), and **B.** a variable loop region (magenta) near the ATP binding site (ATP is in Yellow). Both pictures were made using the BtTLP crystal structure.

TLP/Thg1 Variable Loop Mutants

200 KDa
116.2 KDa
97.4 KDa
66.2 KDa
45 KDa
31 KDa
21.5 KDa

ScThg1
27.8 KDa

Mutant
ScThg1
Should be
~27.8 KDa

BtTLP
28.6 KDa

Mutant BtTLP
~28.6 KDa

DdiTLP4
36 KDa

Mutant
DdiTLP4
Should be
~36 KDa

D68A + D153A ScThg1 Double Mutant

200 KDa
116.2 KDa
97.4 KDa
66.2 KDa
45 KDa
31 KDa
21.5 KDa
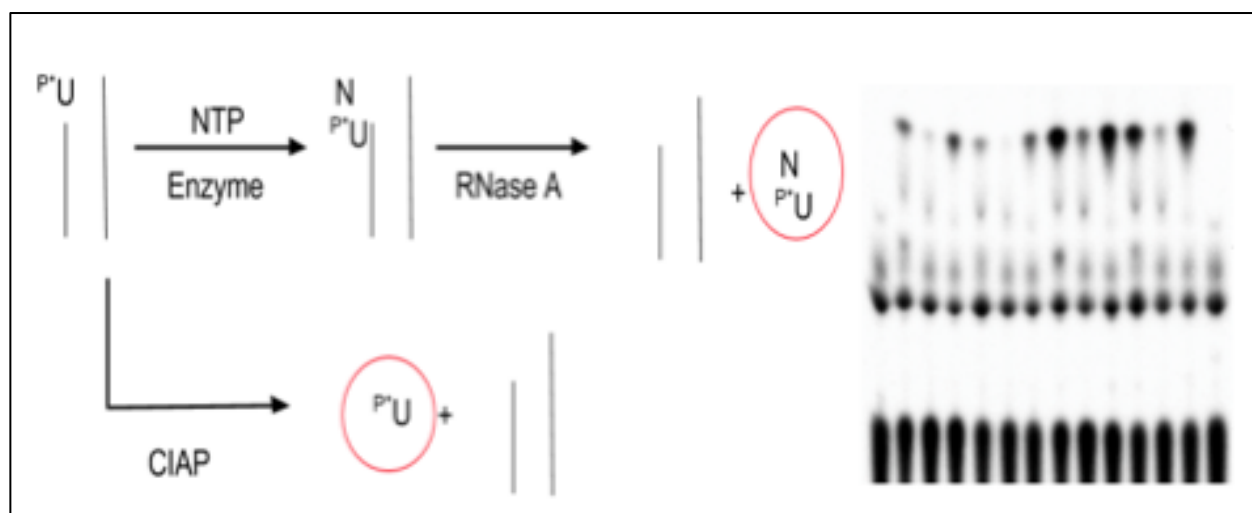
Purified
enzyme
should be
~27.8 KDa

Figure 14: SDS-PAGE gels of the A. single and B, C. double TLP/Thg1 mutants. All mutant enzymes besides the double mutant and the BtTLP variable loop mutant could not be purified.
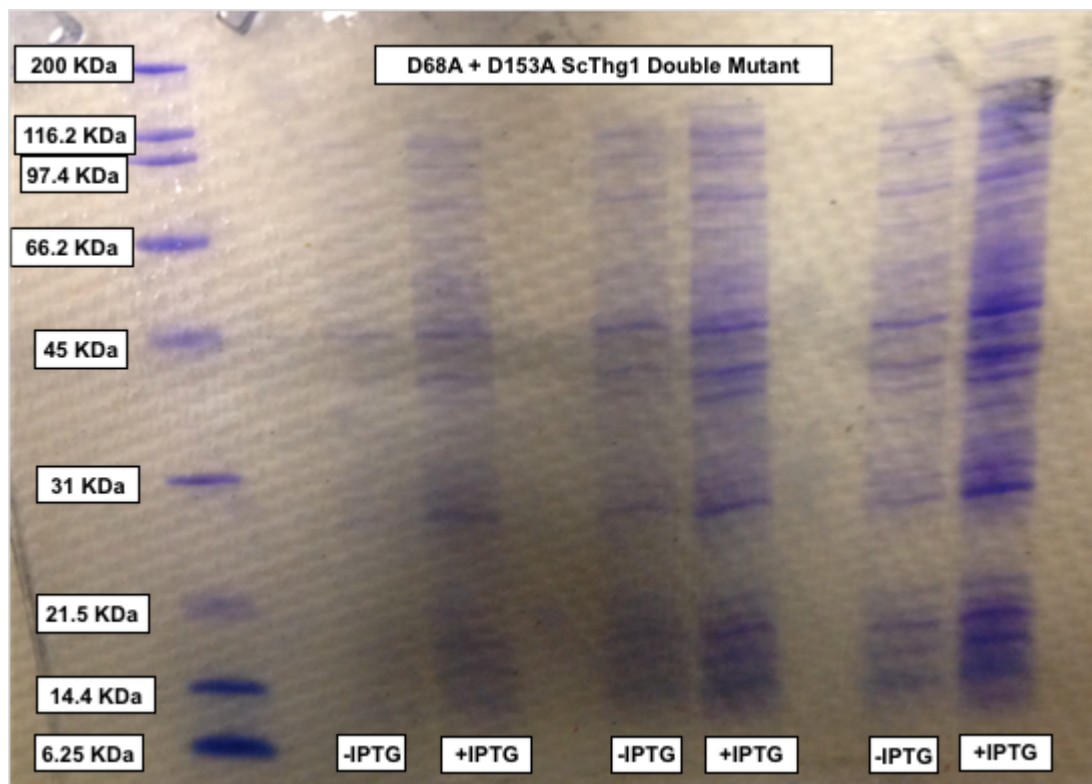
Figure 15: SDS page gel of uninduced and induced D68A + D153A cell cultures. Though 1 mM IPTG was added to the cell culture, the protein failed to overexpress.

Figure 16: Phosphatase Protection TLP activity assay. RNase A cleaves the RNA after the terminal p*U nucleotide. f the reaction occurs, there will be a free $N*p^{32}U$ molecule. If the reaction does not occur, the CIAP will release the *Pi from the unreacted substrate.
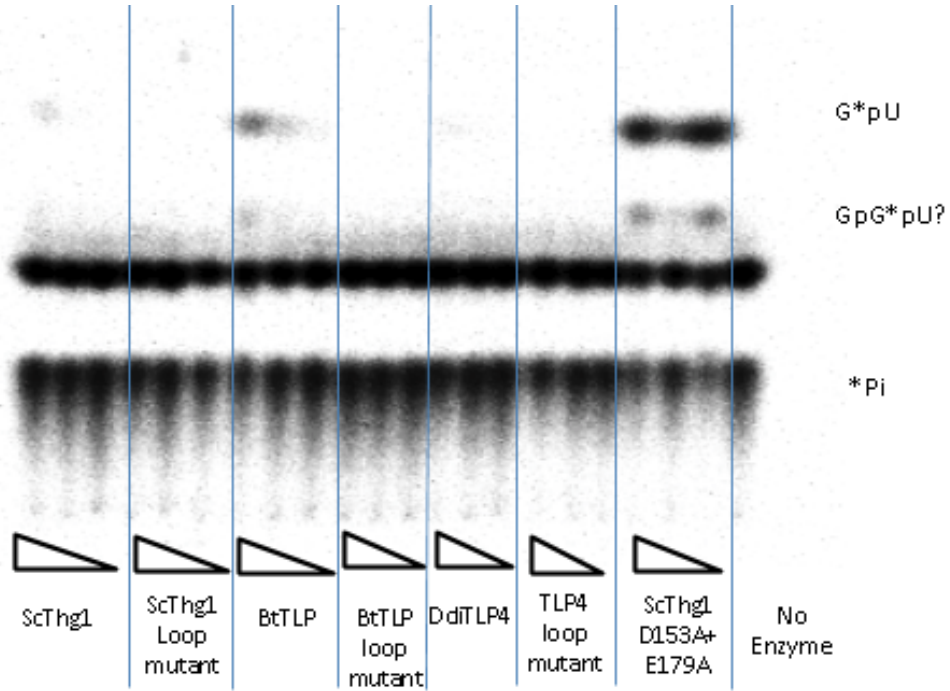
Figure 17: Phosphatase Protection activity assays of the various mutant Thg1/TLP enzymes on 5s rRNA. None of the variable stem loop mutants exhibited any activity. The double D153A + E179A ScThg1 mutant was capable of adding to the 5s rRNA substrate, and seemed to have a higher activity level than BtTLP.

Table 1: Enzyme knockdown at the start of the Vegetative cycle (2 days before developmental start)

Strain	Description	Expected Phenotype	Result
Ddi Thg1 knockdown 1i16	DdiThg1 knockdown	-	Developed more slowly than TLP3, TLP4
Ddi TLP3 knockdown 3i4-11	DdiTLP3 knockdown	-	Developed more quickly than Thg1, V6
Ddi TLP4 knockdown 4i3-15	DdiTLP4 knockdown	-	Developed more quickly than Thg1, V6, and TLP3 Some fruiting bodies were significantly thicker and longer overall
Vector CTL V6	Wild type	WT	WT

Table 2: Enzyme knockdown at the start of the Social cycle

Strain	Description	Expected Phenotype	Result
Ddi Thg1 knockdown 1i16	DdiThg1 knockdown	-	Similar to WT
Ddi TLP3 knockdown 3i4-11	DdiTLP3 knockdown	-	Developed more slowly than Thg1, V6
Ddi TLP4 knockdown 4i3-15	DdiTLP4 knockdown	-	Developed more slowly than Thg1, V6
Vector CTL V6	Wild type	WT	WT

Table 3: Observed rates of nucleotidyl transfer by Thg1, TLPs, onto the minimal RNA substrates

Enzyme	Substrate	1mM* GTP $k_{Obs}(min^{-1})$	1mM* UTP $k_{Obs}(min^{-1})$
D. discoideum TLP **(5uM NTP)**	mini A_{73}	**Not Detectable**	0.15 ± 0.087
	micro A_{73}	**Not Detectable**	0.001
	mini C_{73}	2.61 ± 0.33	**Not Detectable**
	micro C_{73}	0.086 ± 0.024	**Not Detectable**
S. cerevisiae Thg1**	mini A_{73}	0.02	0.009
	micro A_{73}	0.03	0.01
	mini C_{73}	0.14	0.014
	micro C_{73}	1.4	0.008
M. smithii TLP***	mini A_{73}	**Not Detectable**	0.12 ± 0.008
	micro A_{73}	**Not Detectable**	**Not Detectable**
	mini C_{73}	0.022 ± 0.007	0.019 ± 0.008
	micro C_{73}	**Not Detectable**	**Not Detectable**
B. thuringiensis TLP***	mini A_{73}	0.005 ± 0.006	1.16 ± 0.04
	micro A_{73}	**Not Detectable**	**Not Detectable**
	mini C_{73}	0.5 ± 0.2	2.4 ± 1.6
	micro C_{73}	0.006 ± 0.002	0.006 ± 0.002

* 5uM NTP for DdiTLP4
**Data collected by Krishna Patel
***Data collected by Lauren Duff [7]

n (mini A) = 2
n (micro A) = 2
n (mini C) = 2
n (micro C) = 3

Table 4: Mutations made in TLPs, Thg1

Mutation	Reasoning	Outcome
H145Y in AcaTLP2, Y129H point mutations in ScThg1	ncRNA-editing enzymes have a conserved histidine next to the active site	Could not be sequenced
Variable loop domain swaps in BtTLP, DdiTLP4, ScThg1	ncRNA-editing enzymes have conserved nucleotides in this loop, which is somewhat close to the ATP active site	Structure disrupted, enzyme nonfunctional
ScThg1 D68A + D153A double mutant	D68A and D153A point mutations in ScThg1 lead to expanded substrate specificity of the enzyme	Enzyme would not overexpress upon induction
ScThg1 D153A + E179 double mutant	D153A and E179A point mutations in ScThg1 lead to expanded substrate specificity of the enzyme	Enzyme had expanded substrate specificity, enhanced activity levels

Bibliography

[1] Jackman JE, Gott JM, Gray MW. Doing it in reverse: 3'-to-5' polymerization by the Thg1 superfamily. *RNA*. 2012;18(5):886-899. doi:10.1261/rna.032300.112.

[2] Long Y, Jackman JE. Unpublished Data

[3] Aspegren, A. "Novel Non-coding RNAs in Dictyostelium Discoideum and Their Expression during Development." *Nucleic Acids Research* 32.15 (2004): 4646-656. Web.

[4] Abad M.G., Long Y., Willcox A., Gott J.M., Gray M.W., Jackman J.E. *A role for tRNA[His] guanylyltransferase (Thg1)-like proteins from Dictyostelium discoideum in mitochondrial 5'-tRNA editing*. RNA, 2011. 17(4): p. 613-623.

[5] Chisholm RL, Gaudet P, Just EM, Pilcher KE, Fey P, Merchant SN, Kibbe WA. (2006). dictyBase, the model organism database for Dictyostelium discoideum. *Nucleic Acids Res.* 34 (Database issue): D423-7.

[6] Francklyn, C., K. Musier-Forsyth, and P. Schimmel, *Small RNA helices as substrates for aminoacylation and their relationship to charging of transfer RNAs.* European Journal of Biochemistry, 1992. 206(2): p. 315-21.

[7] Duff LE, Jackman JE. Substrate recognition by the tRNA[His]-guanylyltransferase: a kinetic investigation with model RNA substrates. A senior thesis, *The Ohio State University.* April 2015

[8] Choudry SG, Jackman JE. Substrate recognition by the tRNA[His]-guanylyltransferase: a kinetic investigation with model RNA substrates. A senior thesis, *The Ohio State University.* April 2013

[9] Fey, P., Dodson, R., Basu, S., Chisholm, R. L. (2013). 'One Stop Shop for Everything Dictyostelium: dictyBase and the Dicty Stock Center'. **Dictyostelium discoideum Protocols**. *Methods Mol. Biol.* 983:59-92, edited by Ludwig Eichinger and Francisco Rivero.

[10] Hardt, Wolf Dietrich, Judith Schlegl, Volker A. Erdmann, and Roland K. Hartmann. "Role of the D Arm and the Anticodon Arm in TRNA Recognition by Eubacterial and Eukaryotic RNase P Enzymes." *Biochemistry* 32.48 (1993): 13046-3053. Web.

[12] Hori H, Osawa S, Iwabuchi M. The nucleotide sequence of 5S rRNA from a cellular slime mold Dictyostelium discoideum. *Nucleic Acids Research*. 1980;8(23):5535-5539.

[13] Hyde, Samantha J. et al. "tRNAHis Guanylyltransferase (THG1), a Unique 3'-5' Nucleotidyl Transferase, Shares Unexpected Structural Homology with Canonical 5'-3' DNA Polymerases." *Proceedings of the National Academy of Sciences of the United States of America* 107.47 (2010): 20305–20310.